

Fig. 1A

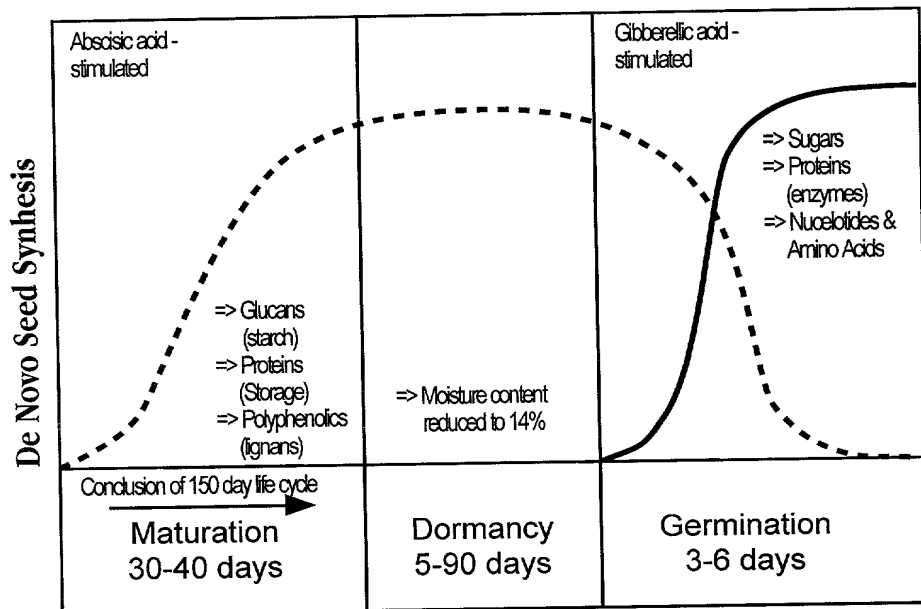


Fig. 1B

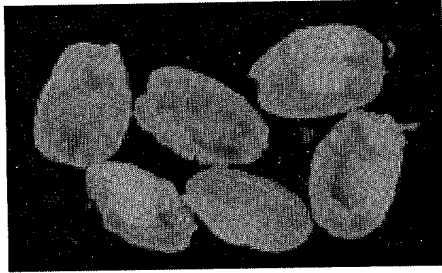


Fig. 1C

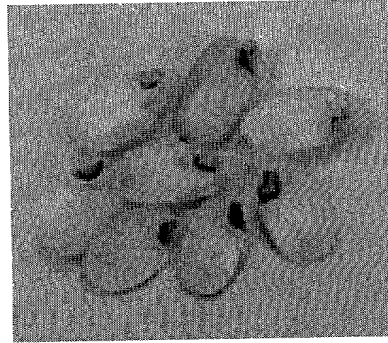


Fig. 1D

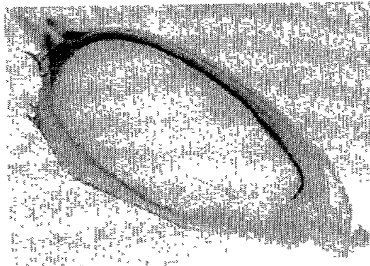


Fig. 1E



Fig. 1F

10	20	30	40	50	60	70
GGTACCCATC	TAATACATTA	ATAACAAGAG	AGAGAATGGA	TAATGCAATT	ATTTATTTTT	ATGGGAGGCT
CCATGGGTAG	ATTATGTAAT	TATTGTTCTC	TCTCTTACCT	ATTACGTAA	TAAATAAAAA	TACCCTCCGA
80	90	100	110	120	130	140
ATATTTTTAT	CGGATTTTAG	TAAATAACGG	GGCAATTCGG	TACTTAGGTA	AAGCTACGTA	TGACTATCGC
TATAAAAATA	GCCTAAAAATC	ATTTATTGCC	CCGTAAAGCC	ATGAATCCAT	TTCGATGCAT	ACTGATAGCG
150	160	170	180	190	200	210
TACCGCTACG	GTAGTTGAAT	TGGAATTCCT	CGATAGCATC	TGTTGTGTTG	TTGCAGTTAG	GGTACTTGAA
ATGGCGATGC	CATCAACTTA	ACCTTAAGAA	GCTATCGTAG	ACAACACAAC	AACGTCAATC	CCATGAACTT
220	230	240	250	260	270	280
TAGCTCCAGC	CGTGAAAACG	AGGGGTTTTTC	GCAGGTTTTTA	TAGGATTGCC	AAGTTAGACT	AGGGCAATTC
ATCGAGGTCG	GCACTTTTGC	TCCCCAAAAG	CGTCCAAAAT	ATCCTAACGG	TTCAATCTGA	TCCCGTTAAG
290	300	310	320	330	340	350
ATGTTACGCG	TATTGTGTAG	TATATGAAAA	AGGAGATCTC	CCAAACAATT	TATAATTTTG	TATAAGGGAG
TACAAGTGCC	ATAACACATC	ATATACTTTT	TCCTCTAGAG	GGTTTGTAA	ATATTAAAAC	ATATTCCTC
				>AT-rich_region_		
360	370	380	390	400	410	420
AAATCGAACT	TGAGGTGTCT	AATTCACCAA	CCGAGCTACT	CCCTCCGTTT	CATATATGTA	TATACATATA
TTTAGCTTGA	ACTCCACAGA	TTAAGTGGTT	GGCTCGATGA	GGGAGGCAAA	GTATATACAT	ATATGTATAT
430	440	450	460	470	480	490
TACGTATATA	TACGTATATA	CACATATACG	TATATACATA	TATGGTATAT	ACATATATAT	ATATATATAT
ATGCATATAT	ATGCATATAT	GTGTATATGC	ATATATGTAT	ATACCATATA	TGTATATATA	TATATATATA
500	510	520	530	540	550	560
ATATATATAT	ATGTGTGTGT	GTGTATGTGG	GGTGGCAATG	CTAAAAAGTT	TTATAATATG	AACGGATGAA
TATATATATA	TACACACACA	CACATACACC	CCACCGTTAC	GATTTTTCAA	AATATTATAC	TTGCCTACTT
570	580	590	600	610	620	630
GTACTATCCA	CTAAGTCCCT	ATAGTTTTCT	GGCACTGTGT	AGTATACGAA	TGCACAATTA	TATCCATAAA
CATGATAGGT	GATTCAGGGA	TATCAAAAGA	CCGTGACACA	TCATATGCTT	ACGTGTTAAT	ATAGGTATTT
640	650	660	670	680	690	700
ATTGATATTA	TATATTCGTC	GCGACGAAAA	TAAAGACATA	ATATTCGGTA	TACCATTTAT	CCACGATATA
TAACTATAAT	ATATAAGCAG	CGCTGCTTTT	ATTTCTGTAT	TATAAGCCAT	ATGGTAAATA	GGTGCTATAT
710	720	730	740	750	760	770
TCTAAATTCC	ACTGATATAT	CTAAATTCCA	CTTGATCCCT	TTTATGGATA	AATTCTGGAT	AACAATTACT
AGATTTAAGG	TGACTATATA	GATTTAAGGT	GAAGTAGGGA	AAATACCTAT	TTAAGACCTA	TTGTTAATGA
780	790	800	810	820	830	840
ACCAGCAGTA	TATCCTACTA	TCAGCGCACT	GCACACCAAA	CTACCCTCAC	CCAGTAGTTA	CAAACGCATA
TGGTCGTCAT	ATAGGATGAT	AGTCGCGTGA	CGTGTGGTTT	GATGGGAGTG	GGTCATCAAT	GTTTGCCTAT

Fig. 2A

```

      850      860      870      880      890      900      910
TTTTGCCGTT AGTTAATTAT TATCCGGTAA AGAAGGTAAA GAAGATTGGT AGTAATCCAA AATTTTCCCA
AAAACGGCAA TCAATTAATA ATAGGCCATT TCTTCCATTT CTTCTAACCA TCATTAGGTT TTAAAAGGGT

      920      930      940      950      960      970      980
ACCCCAACCT CGGAACAAAA ACCGCGTAGT ATTTGTCGTA ACCAGGAGCA TCCGAGTCAT TAATTTACAC
TGGGGTTGGA GCCTTGTTTT TGGCGCATCA TAAACAGCAT TGGTCCTCGT AGGCTCAGTA ATTAAATGTG

      >Transcription_start_site
>CAAG_site

      990      1000      1010      1020      1030      1040      1050
CCAAACACAA AAAATTAGCA GCACGCAGCC GCCTTCCCAA TCCTCTCCTC TCTCCTCTCC TCTTCTCCAA
GGTTTGTTGTT TTTTAATCGT CGTGCGTCGG CGGAAGGGTT AGGAGAGGAG AGAGGAGAGG AGAAGAGGTT

      1060      1070      1080      1090      1100      1110      1120
GCGGCAATTC GCGCGAGGTT TTCTCCGATC AAACCCTCGA ATCCCCCCTT CGCGAATCCA TCGGAGGGTA
CGCCGTTAAG CGCGCTCCAA AAGAGGCTAG TTTGGGAGCT TAGGGGGGGA GCGCTTAGGT AGCCTCCCAT

      1130      1140      1150      1160      1170      1180
GCCCCGCGAT CCGCGTCGGC GAGAGCGGAT TCCGATTCCG CG ATG GAG CGG GTG TTC TCC GTG
CGGGGCGCTA GGCGCAGCCG CTCTCGCCTA AGGCTAAGGC GC TAC CTC GCC CAC AAG AGG CAC
      M E R V F S V>
      _a_a_EXON1_a_a_>

      1190      1200      1210      1220      1230      1240
GAG GAG ATC TCC GAC CCA TTC TGG GTC CCG CCT CCG CCG CCG CAG TCG GCG GCG GCG
CTC CTC TAG AGG CTG GGT AAG ACC CAG GGC GGA GGC GGC GGC GTC AGC CGC CGC CGC
E E I S D P F W V P P P P P Q S A A A>
_a_a_a_a_a_a_a_a_EXON1_a_a_a_a_a_a_a_a_a_a_>

      1250      1260      1270      1280      1290
GCC CAG CAG CAG GGC GGC GGC GGC GTG GCT TCG GGA GGT GGT GGT GGT GTA GCG GGG
CGG GTC GTC GTC CCG CCG CCG CCG CAC CGA AGC CCT CCA CCA CCA CCA CAT CGC CCC
A Q Q Q G G G G V A S G G G G G V A G>
_a_a_a_a_a_a_a_a_EXON1_a_a_a_a_a_a_a_a_a_a_>

1300      1310      1320      1330      1340      1350
GGC GGC GGC GGC GGC AAC GCG ATG AAC CGG TGC CCG TCG GAG TGG TAC TTC CAG AAG
CCG CCG CCG CCG CCC TTG CGC TAC TTG GCC ACG GGC AGC CTC ACC ATG AAG GTC TTC
G G G G G N A M N R C P S E W Y F Q K>
_a_a_a_a_a_a_a_a_EXON1_a_a_a_a_a_a_a_a_a_a_>

      1360      1370      1380      1390      1400      1410
TTT CTG GAG GAG GCG GTG CTC GAT AGC CCC GTC CCG AAC CCT AGC CCG AGG GCC GAA
AAA GAC CTC CTC CGC CAC GAG CTA TCG GGG CAG GGC TTG GGA TCG GGC TCC CGG CTT
F L E E A V L D S P V P N P S P R A E>
_a_a_a_a_a_a_a_a_EXON1_a_a_a_a_a_a_a_a_a_a_>

```

Fig. 2B

2080 2090 2100 2110 2120 2130 2140
 CTTTAGCACC TCTTATCTTA TCAACCATGG TGAATAAATT GAAGGGGGGA CTCAGGGGGG TATCCATGGG
 GAAATCGTGG AGAATAGAAT AGTTGGTACC ACTTTTTTTAA CTTCCCCCCT GAGTCCCCCC ATAGGTACCC

2150 2160 2170 2180 2190 2200 2210
 TCCGATGGGT GCAGGGGGGA CTGAGTCCCC CCTGCACCCA CGTTGAATCC GCCCTGGCAT GCGTATAAGC
 AGGCTACCCA CGTCCCCCCT GACTCAGGGG GGACGTGGGT GCAACTTAGG CGGGACCGTA CGCATATTCG

2220 2230 2240 2250 2260 2270 2280
 TGTCACAGCC ATTTCTAGGT GCTTGTGCTT AGTTGGGTGA TGTCAGCTTA ATTTGTCTTT TCTATGTCGT
 ACAGTGTCGG TAAAGATCCA CGAACACGAA TCAACCCACT ACAGTCGAAT TAAACAGAAA AGATACAGCA

2290 2300 2310 2320 2330 2340 2350
 CATCGATTTT CTAAGAAACG AAAAATAGCC TATTTATGTG CTCCAGAATT TGATGATCCC TGGCCCTTCA
 GTAGCTAAAA GATTCTTTGC TTTTATATCGG ATAAATACAC GAGGTCTTAA ACTACTAGGG ACCGGGAAGT

2360 2370 2380 2390 2400 2410 2420
 TTTGCTGAAA TTAGCCTATT TGTTGGTTGC CCTTCAGTTT TTTCCCAGCT TATGTTGTTG CAATGTGTGG
 AAACGACTTT AATCGGATAA ACAACCAACG GGAAGTCAAA AAAGGGTCGA ATACAACAAC GTTACACACC

2430 2440 2450 2460 2470 2480 2490
 CTATGCCTCG TTTTGTGCCC TATAATTTAT TATTTGCAAT TCATTTTTGT ACATGACTTA AAATGACACT
 GATACGGAGC AAAACACGGG ATATTAAATA ATAAACGTTA AGTAAAAACA TGTACTGAAT TTTACTGTGA

2500 2510 2520 2530 2540 2550 2560
 AGAGCAACAT GCACTGATTG GTTATCCTAT AATCATTTAT GTAGTTCTGT TCATTTTATC ATGCTAGCTC
 TCTCGTTGTA CGTGACTAAC CAATAGGATA TTAGTAAATA CATCAAGACA AGTAAAATAG TACGATCGAG

2570 2580 2590 2600 2610 2620
 ATGTCATTTT CATCTTCAG GCC TCT GGC ACA GTT CCA CCT GAG CGT CCT GGA GCT GGT TCA
 TACAGTAAAA GTAGAAGTC CGG AGA CCG TGT CAA GGT GGA CTC GCA GGA CCT CGA CCA AGT
 A S G T V P P E R P G A G S>
 _b_b_b_b_b_b_b_EXON2_b_b_b_b_b_b_b_>

2630 2640 2650 2660 2670 2680
 TCC TTG CTG AAT GCA GAT GTT TCA CAC ATA GGC GCT CCT AAT TCC ATC GGA GTTACTTA
 AGG AAC GAC TTA CGT CTA CAA AGT GTG TAT CCG CGA GGA TTA AGG TAG CCT CCATGAAT
 S L L N A D V S H I G A P N S I G>
 _b_b_b_b_b_b_b_EXON2_b_b_b_b_b_b_b_>

2690 2700 2710 2720 2730 2740 2750
 TCTTATCTGG TTACATTTTC AGATTGTTAT GAAACTACCC AAATATCCTG CACAATTGCA TGGGATTAAA
 AGAATAGACC AATGTAAAAG TCTAACAATA CTTTGATGGG TTTATAGGAC GTGTTAACGT ACCCTAATTT

Fig. 2D

3410	3420	3430	3440	3450	3460	3470
AGGTGATC	ATTCATTGCT	TCCTTGTAAT	ATAGATTCTG	TACATAATTA	ACCTACCTCG	TCATGCATGC
TCCACTAG	TAAGTAACGA	AGGAACATTA	TATCTAAGAC	ATGTATTAAT	TGGATGGAGC	AGTACGTACG
3480	3490	3500	3510	3520	3530	3540
ATGTGTCCTA	TTTTTCACCTT	AGCCCTTTCA	GTTGGATTTT	CACTTTTCATC	CGGTAGCCTT	TCAGTTTCCT
TACACAGGAT	AAAAGTGGAA	TCGGGAAAGT	CAACCTAAAG	GTGAAAGTAG	GCCATCGGAA	AGTCAAAGGA
3550	3560	3570	3580	3590	3600	3610
ATTGCATCGC	ATATATGATC	TTTTTACCTAC	CATATTAGTT	CTCTGTGTGC	CATACTCAGT	GCTTAGTGTC
TAACGTAGCG	TATATACTAG	AAAATGGATG	GTATAATCAA	GAGACACACG	GTATGAGTCA	CGAATCACAG
3620	3630	3640	3650	3660	3670	3680
TCGAGCAAGA	GAGGAATTTG	TATGGCTATT	ACACGTAGCA	CTTTGCTCTC	TACTTGTTTA	TTGACATAAG
AGCTCGTTCT	CTCCTTAAAC	ATACCGATAA	TGTGCATCGT	GAAACGAGAG	ATGAACAAAT	AACTGTATTC
3690	3700	3710	3720	3730	3740	3750
CAATTTGGGA	TGAATTAAAT	CTGAGTTCAC	ATCATATTCC	TTATGTCACA	AGTTTCTGAA	ACCGATTGTA
GTTAAACCCT	ACTTAATTTA	GACTCAAGTG	TAGTATAAGG	AATACAGTGT	TCAAAGACTT	TGGCTAACAT
3760	3770	3780	3790	3800	3810	3820
TCTAGTATCT	GGTTGATGCA	CCCCCATCTT	GGATTTGCAA	ATCAAAGTTA	TACTCCCTAG	AGAGCTTTAC
AGATCATAGA	CCAACTACGT	GGGGGTAGAA	CCTAAACGTT	TAGTTTCAAT	ATGAGGGATC	TCTCGAAATG
3830	3840	3850	3860	3870	3880	3890
CTTTCATAAA	GCAATTACCC	CAATAAACCA	CGGATTTGAT	AGCTATTGAC	TATGATTACC	AGAATTCATT
GAAAGTATTT	CGTTAATGGG	GTTATTTGGT	GCCTAAACTA	TCGATAACTG	ATACTAATGG	TCTTAAGTAA
3900	3910	3920	3930	3940	3950	3960
TGGCAGCTAT	TTTCTCAATT	TAAGTTTGGT	ATTAGTCTCA	GTTGGCTGTA	AAATAATGTC	ACGGTAGGGT
ACCGTCGATA	AAAGAGTTAA	ATTCAAACCA	TAATCAGAGT	CAACCGACAT	TTTATTACAG	TGCCATCCCA
3970	3980	3990	4000	4010	4020	4030
ACATGTATGT	GCAGCATACA	AGGTATGGGT	GAGTTATGAT	ATGGACAGTG	TGTACACCCC	ACATTTGCTC
TGTACATACA	CGTCGTATGT	TCCATACCCA	CTCAATACTA	TACCTGTCAC	ACATGTGGGG	TGTAAACGAG
4040	4050	4060	4070	4080	4090	4100
ACTAAATCA	AAATATTCAA	ACGTCACGTG	ATGATATGGT	GGATTGCATT	ATACCTTGTA	TTGTTTATTA
TGATTTTAGT	TTTATAAGTT	TGCAGTGCAC	TACTATACCA	CCTAACGTAA	TATGGAACAT	AACAAATAAT
4110	4120	4130	4140	4150	4160	4170
TGTTACTTGT	GCTAGACAAT	AATATAGGCT	GTTCTTTTGG	GTGATTTTGT	ATGAAGATGT	TGAGCAAGCA
ACAATGAACA	CGATCTGTTA	TTATATCCGA	CAAGAAAACC	CACTAAAACA	TACTTCTACA	ACTCGTTCGT
4180	4190	4200	4210	4220	4230	
CTTCTCGATA	TAATGCTAGT	TTTGTTGACC	TGTTCC	AGG AAG CAA TCC AAT CGG GAG TCA GCC		
GAAGAGCTAT	ATTACGATCA	AAACAACCTGG	ACAAGG TCC TTC GTT AGG TTA GCC CTC AGT CGG			
				R K Q S N R E S A>		
				___d___d___d___EXON4___d___d___d___>		

Fig. 2F

5370	5380	5390	5400	5410	5420	5430
AATATTTTGA	TCGATGCTTC	CTCTTGCTCT	TTGCTCTTAA	GCAACCAAGC	ATAAAGATAT	CACTACCTTT
TTATAAAACT	AGCTACGAAG	GAGAACAGAA	AACGAGAATT	CGTTGGTTTC	TATTTCTATA	GTGATGGAAA
5440	5450	5460	5470	5480	5490	5500
TGAGCTGTTT	ATTTGAAGTG	CAAAGCTAAG	CTCAATATCT	CAGGTGTTCA	TTTGAAGTTT	AAAGGTGAAC
ACTCGACAAG	TAAACTTCAC	GTTTCGATT	GAGTTATAGA	GTCCACAAGT	AAACTTCAAA	TTTCCACTTG
5510	5520	5530	5540	5550	5560	5570
TGATAACAAA	CGTCAGGCTA	TGGTGAATGA	AGGGACGTGT	ACATCCCTAA	TACATGTCAT	TTTCATAATC
ACTATTGTTT	GCAGTCCGAT	ACCACTTACT	TCCCTGCACA	TGTAGGGATT	ATGTACAGTA	AAAGTATTAG
5580	5590	5600	5610	5620	5630	5640
AAATTAGTTG	ATGCATTTTC	ACCCAGAATC	CCATCACAGT	TCATCATACA	AGCAAGTGTA	GTTATTAATG
TTTAATCAAC	TACGTAAAG	TGGGTCTTAG	GGTAGTGTCA	AGTAGTATGT	TCGTTACAT	CAATAATTAC
5650	5660	5670	5680	5690	5700	5710
GTAAATTTTT	CGTTTAGAGA	AAAAAAAAGG	AAGCCTTATA	TAAGATTCAC	CGGTGGGGTG	TGAACAATAA
CATTTAAAAA	GCAAATCTCT	TTTTTTTTCC	TTCGGAATAT	ATTCTAAGTG	GCCACCCAC	ACTTGTTATT
5720	5730	5740	5750	5760	5770	5780
TCAATGAATG	AGATCGCATC	CCGTAAGGGC	AGCCTAGCTA	GACAAAAATG	CATAAACTC	CGTATACCAA
AGTTACTTAC	TCTAGCGTAG	GGCATTCCTG	TCGGATCGAT	CTGTTTTTAC	GTATTTTGAG	GCATATGGTT
5790	5800	5810	5820	5830	5840	5850
CCACAACAAC	GCTTGCGCAC	GCGCTCAAAT	GGCAGCGACT	TCATCGCTTT	CGCGGGCAAG	AAACGAATCA
GGTGTTGTTG	CGAACGCGTG	CGCGAGTTTA	CCGTCGCTGA	AGTAGCGAAA	GCGCCCGTTC	TTTGCTTAGT
5860	5870	5880	5890	5900	5910	5920
AGTGATACAT	TGGCAGGGAA	CCACCAAAAG	AAGGCCATCC	AATCCAATCC	ACTCCAACGC	GGCATGGAAG
TCACTATGTA	ACCGTCCCTT	GGTGTTTTTC	TTCCGGTAGG	TTAGGTTAGG	TGAGGTTGCG	CCGTACCTTC
5930	5940	5950	5960	5970	5980	5990
ACAAGACAGA	TGATTCACAG	CTATCTTCTG	CTTCTACAAG	TTTGATACTT	TGTACTGTCC	TTTCAGGGAA
TGTTCTGTCT	ACTAAGTGTC	GATAGAAGAC	GAAGATGTTC	AACTATGAA	ACATGACAGG	AAAGTCCCTT
6000	6010	6020	6030	6040	6050	6060
AAAAGAGCAT	CAGATTAGTC	TGATCTCGGG	CGCGTTGAGT	TCTTGTTGGA	GATCTTGTTG	TGGAGTGGCA
TTTTCTCGTA	GTCTAATCAG	ACTAGAGCCC	GCGCAACTCA	AGAACACCCT	CTAGAACAAC	ACCTCACCGT
6070	6080	6090	6100	6110	6120	6130
GGAGTGACGA	TCGGCTGCCC	CGTTTTCTTC	TACCGAAACA	TCGCCAGTAA	AGAAGCCAAA	AAGACAATAA
CCTCACTGCT	AGCCGACGGG	GCAAAGAAG	ATGGCTTTGT	AGCGGTCATT	TCTTCGGTTT	TTCTGTTATT
6140	6150	6160	6170	6180	6190	6200
TACGGCAATG	GGGATCGCCC	ATCTGCATAA	AACATTGCAT	GACGGAACTG	ATTAATACAA	GAATGACATG
ATGCCGTTAC	CCCTAGCGGG	TAGACGTATT	TTGTAACGTA	CTGCCTTGAC	TAATTATGTT	CTTACTGTAC
6210	6220					
TAAGCTGATA	ATTACGCGTG	CAAGCTT				
ATTCGACTAT	TAATGCGCAC	GTTCGAA				

Fig. 2I

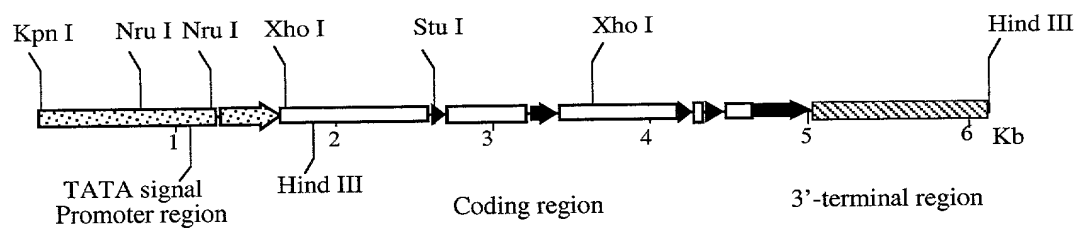


Fig. 3

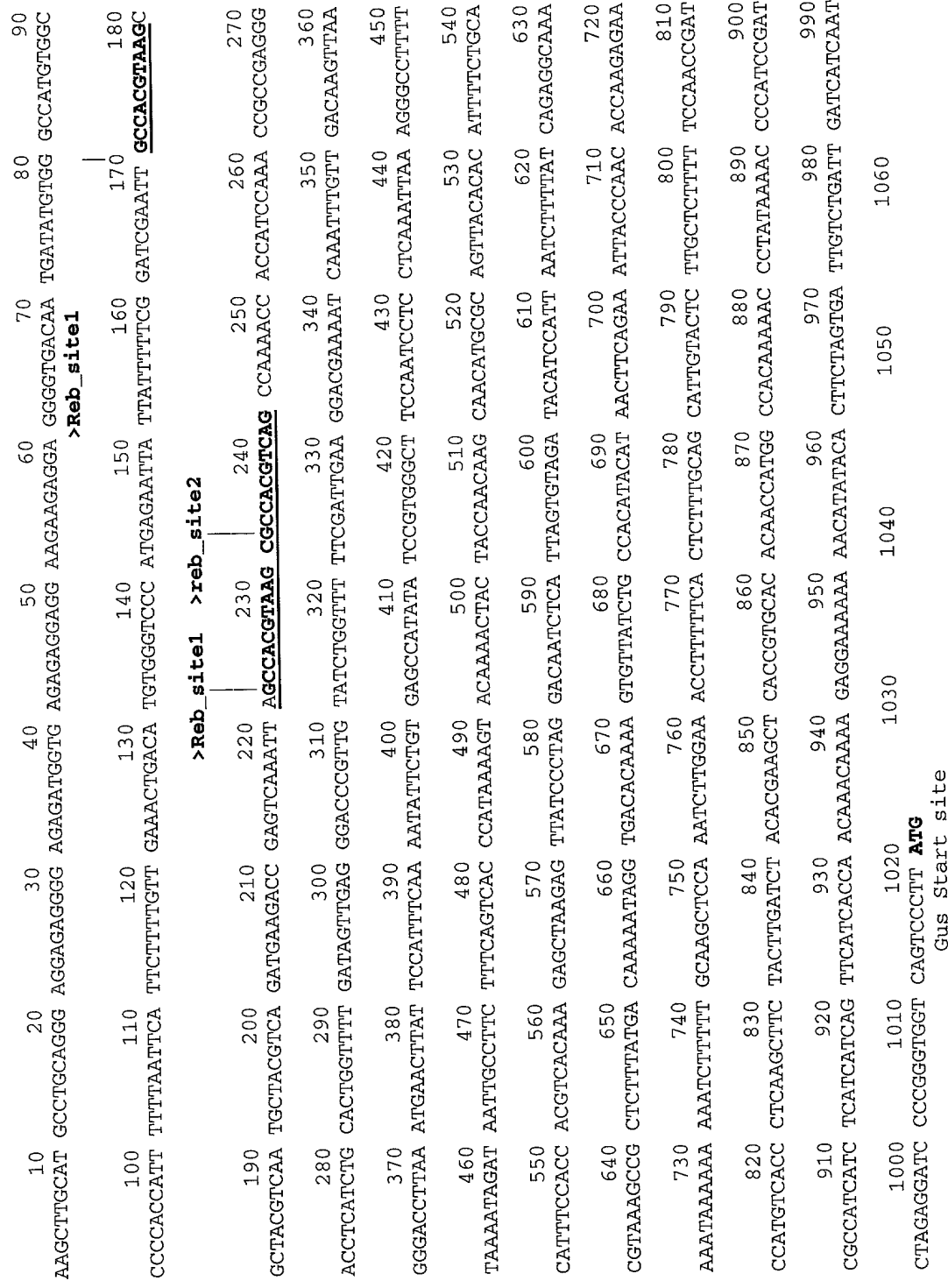


Fig. 4

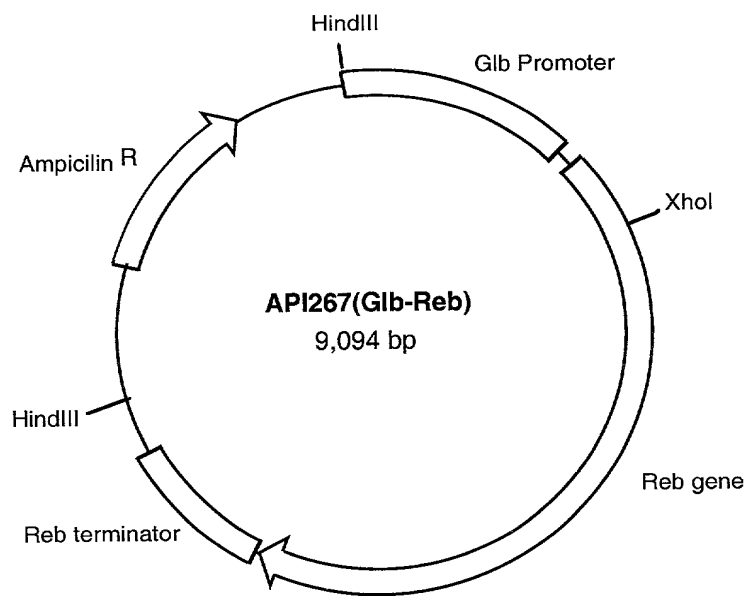


Fig. 5A

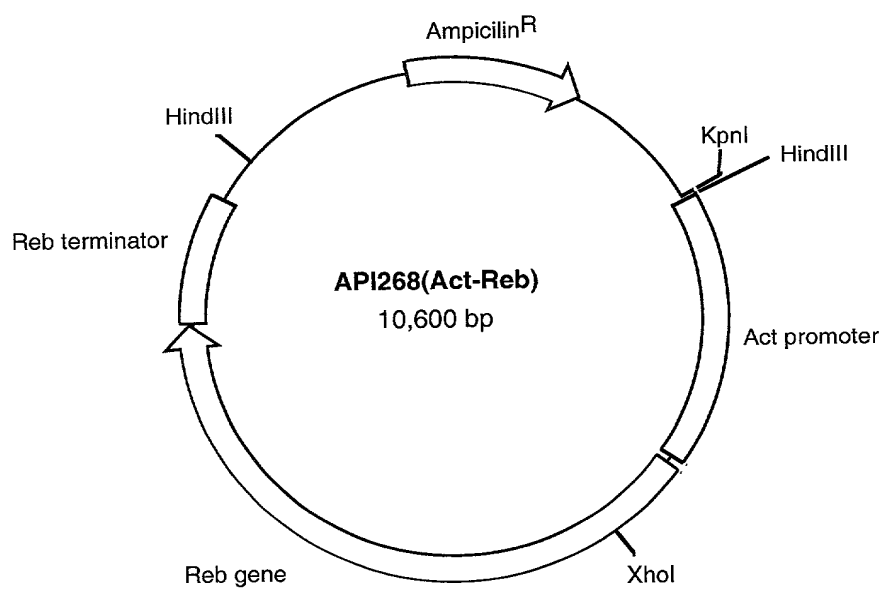


Fig. 5B

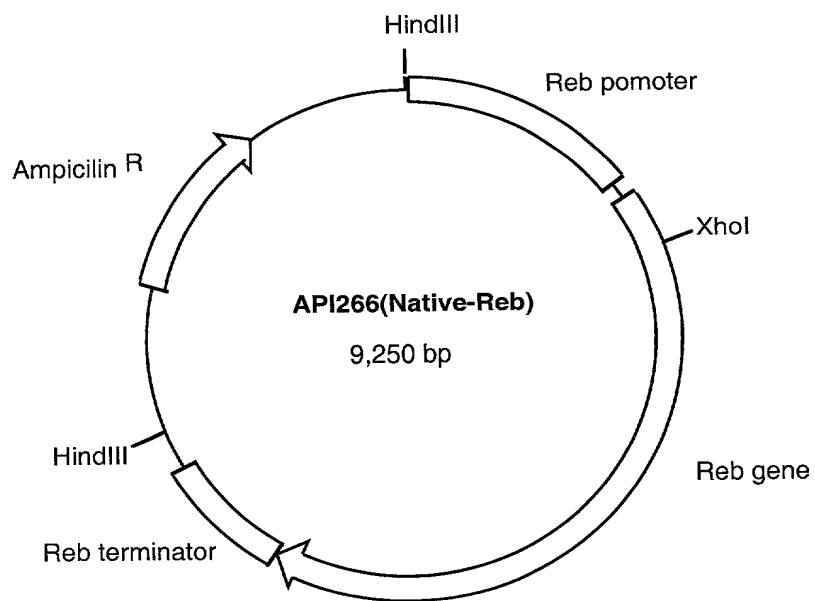


Fig. 5C

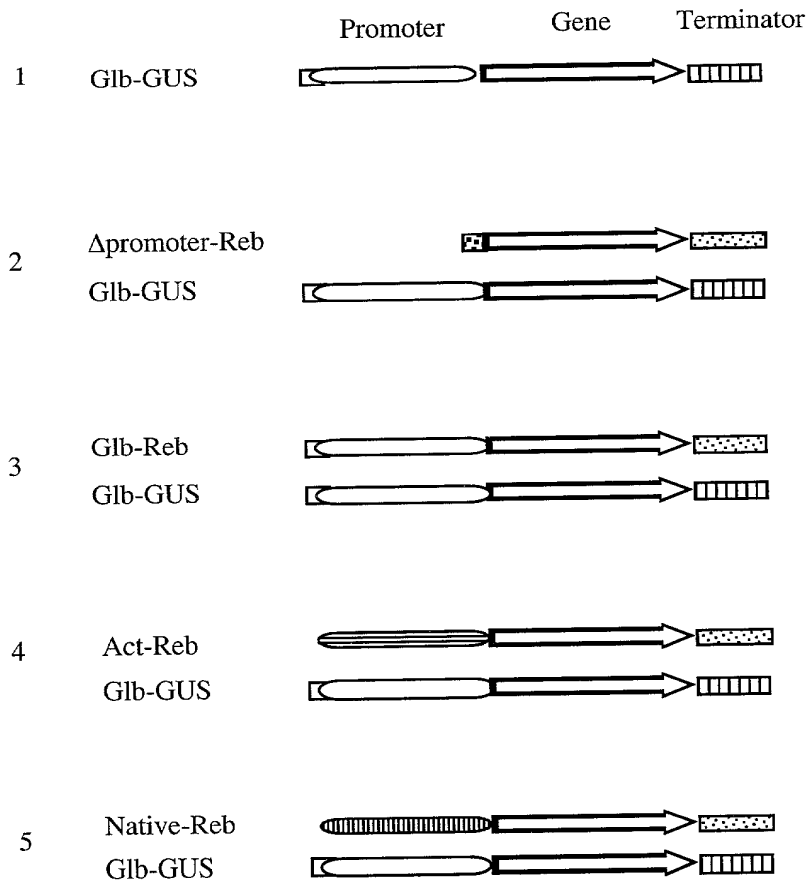


Fig. 6A

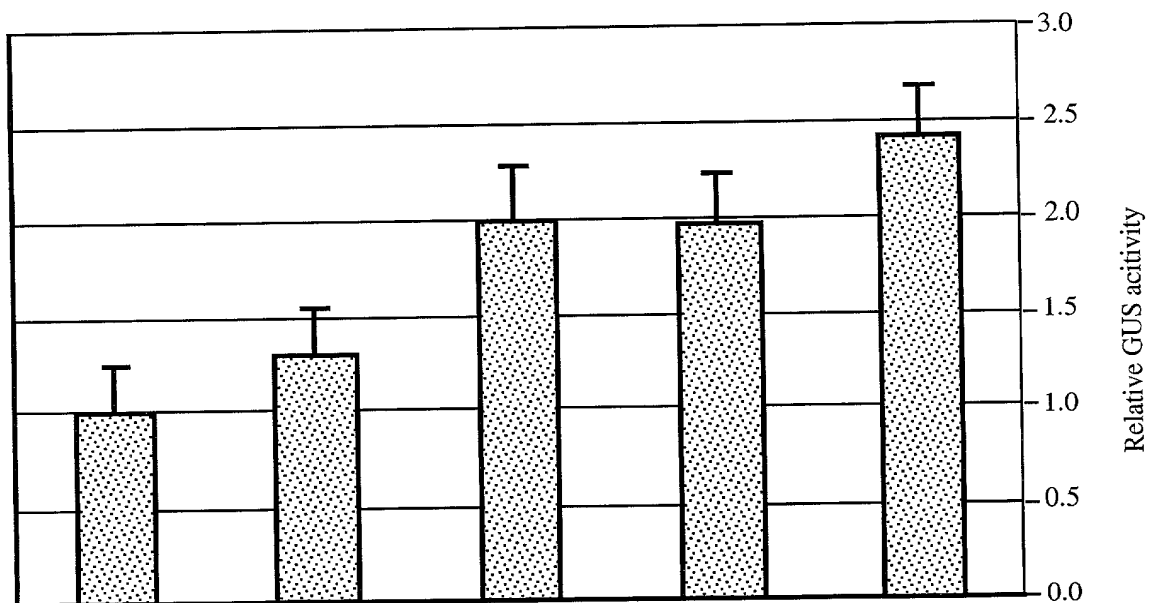


Fig. 6B

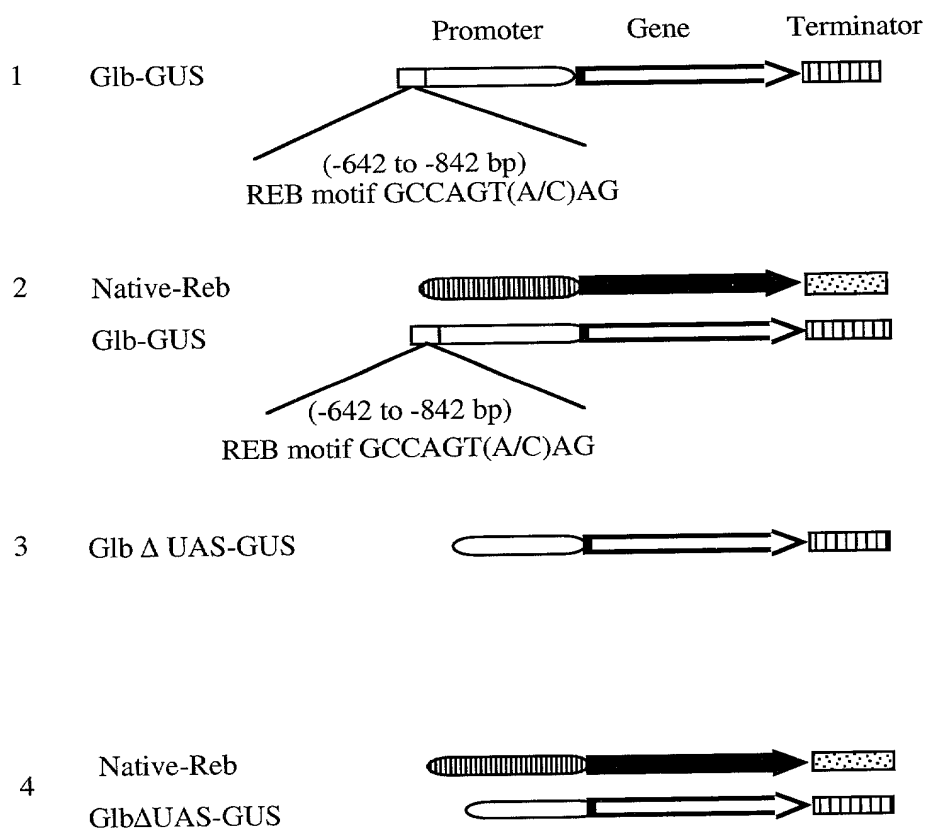


Fig. 7A

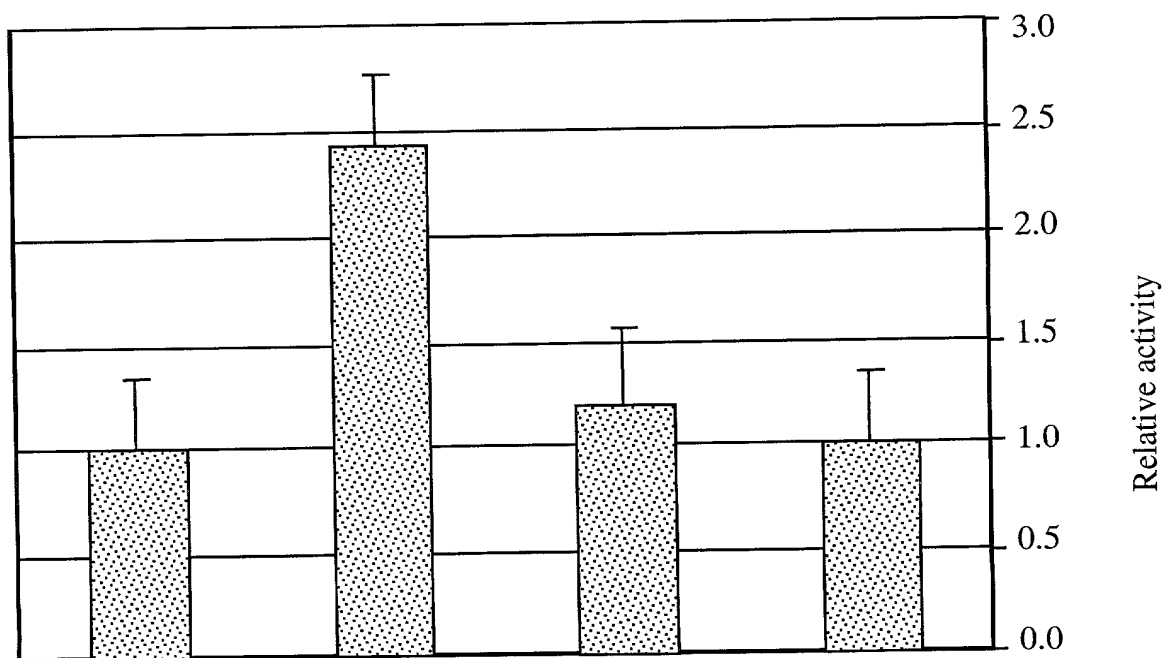


Fig. 7B

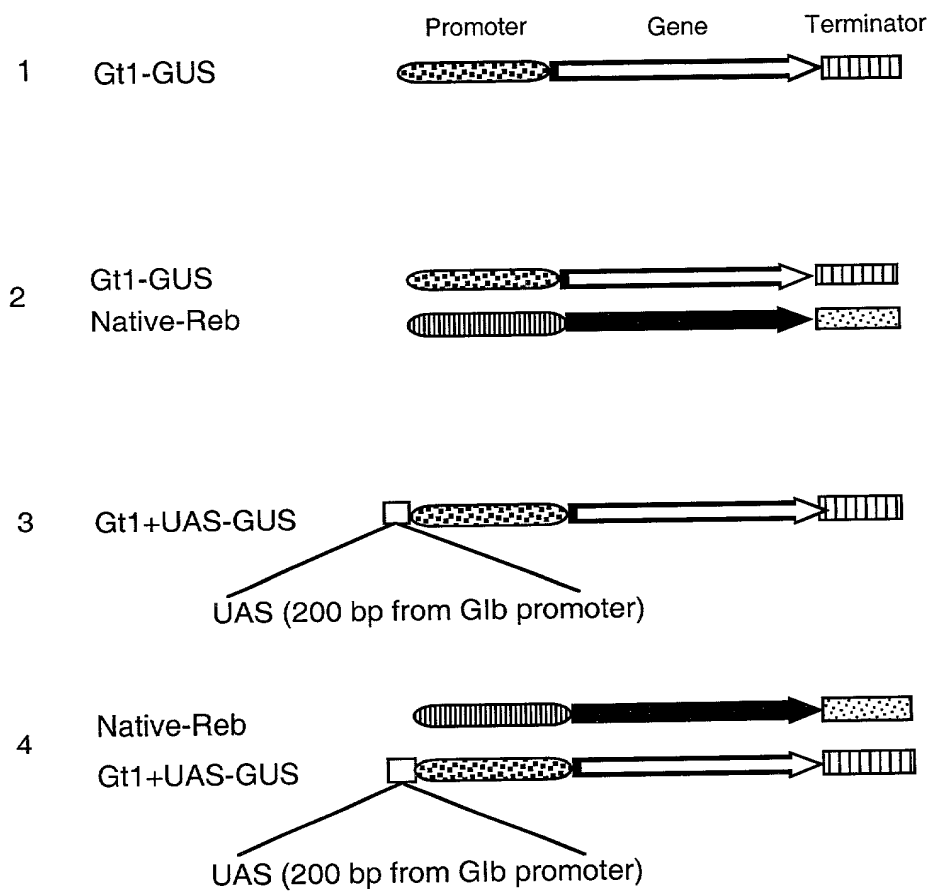


Fig. 8A

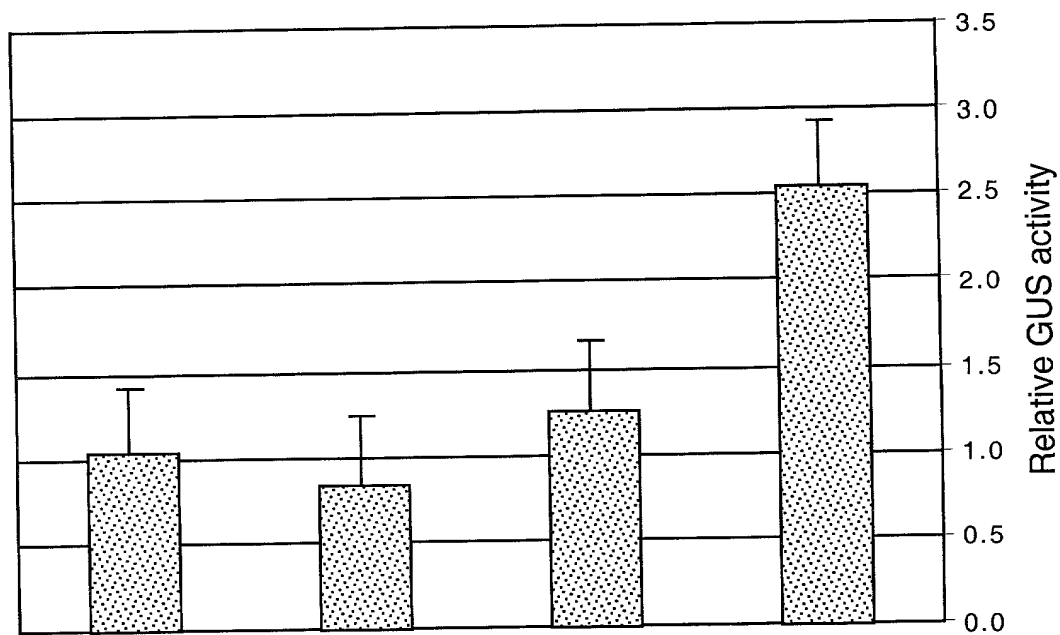


Fig. 8B

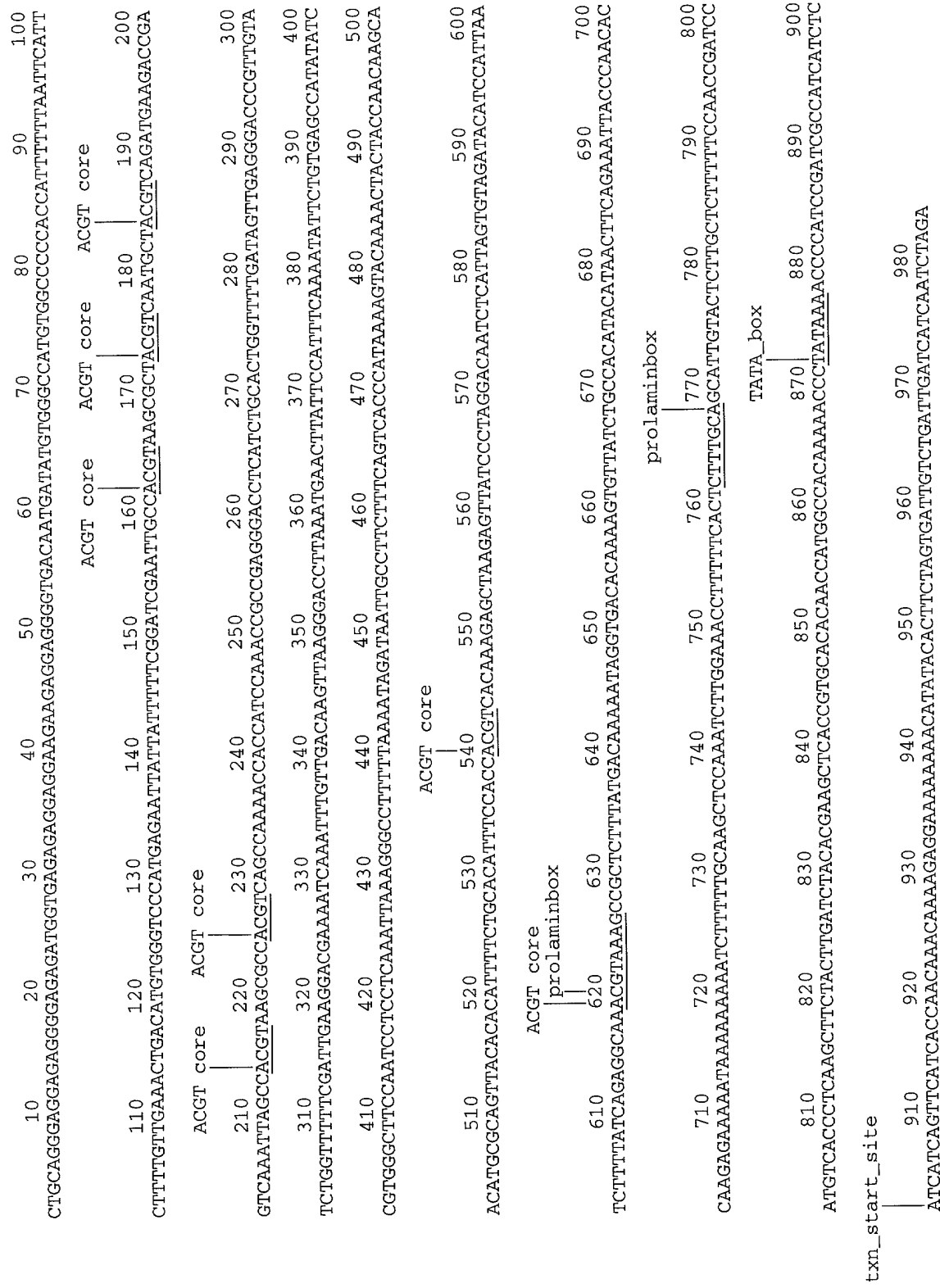


Fig. 9

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

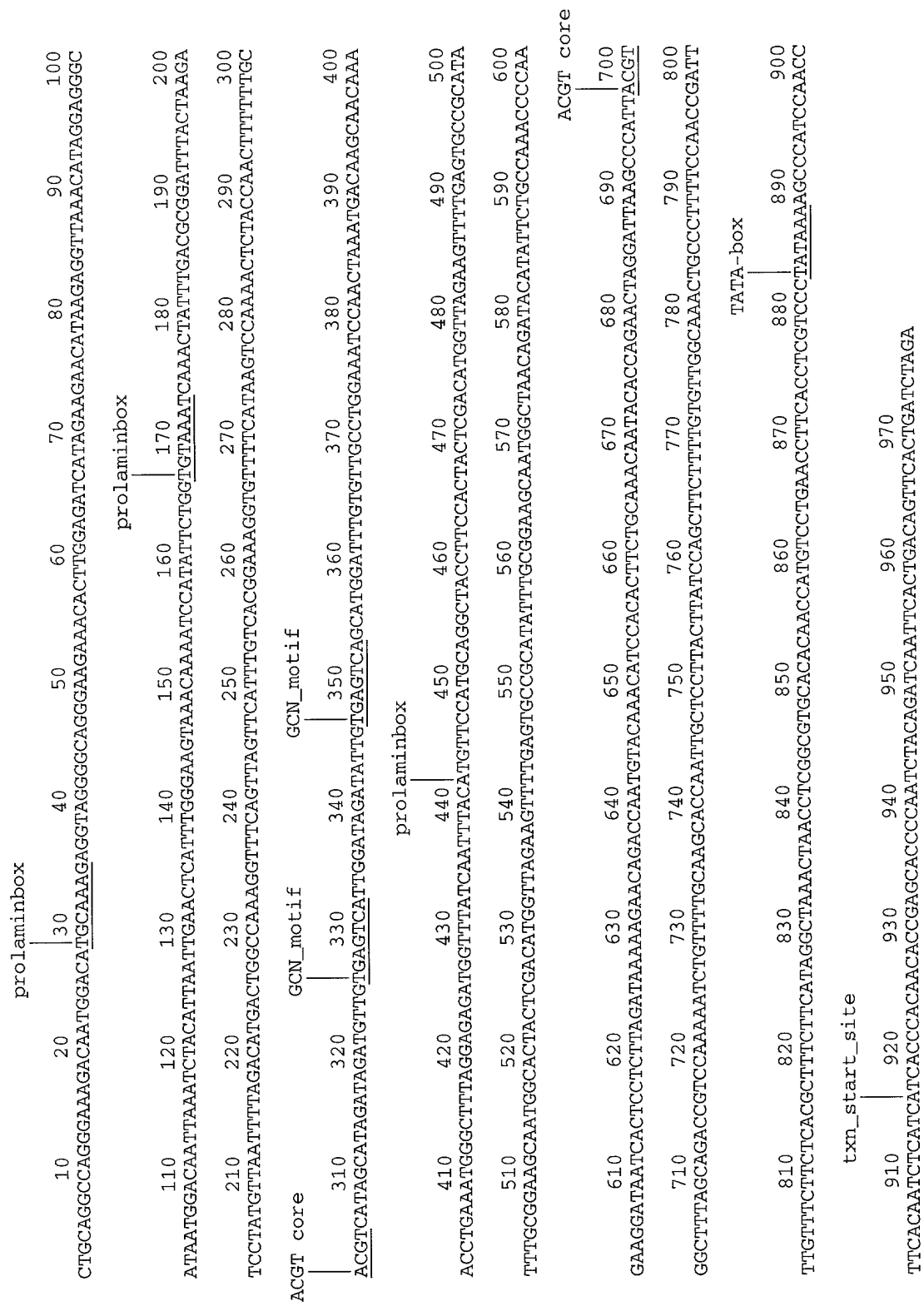


Fig. 10

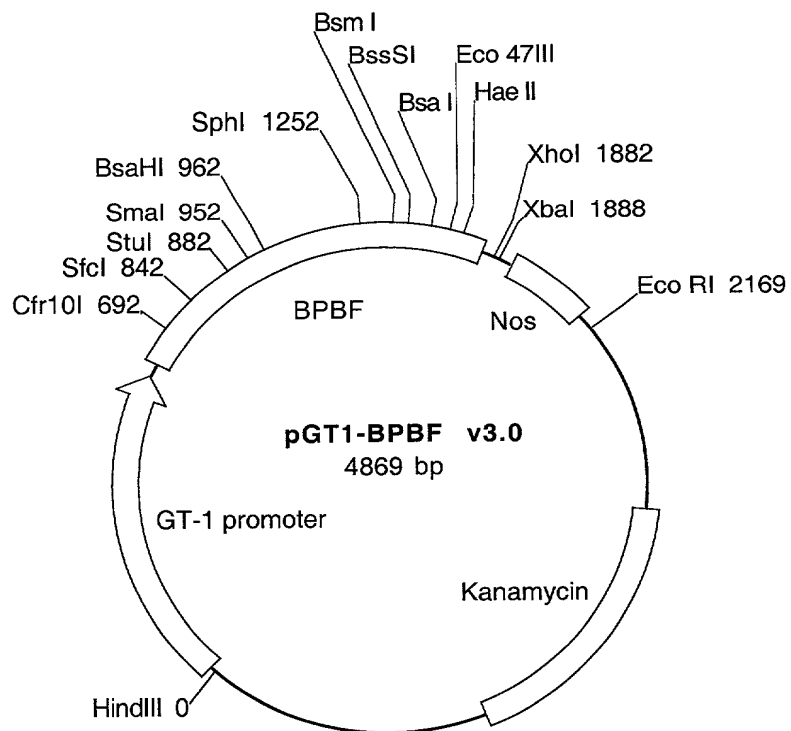


Fig. 11A

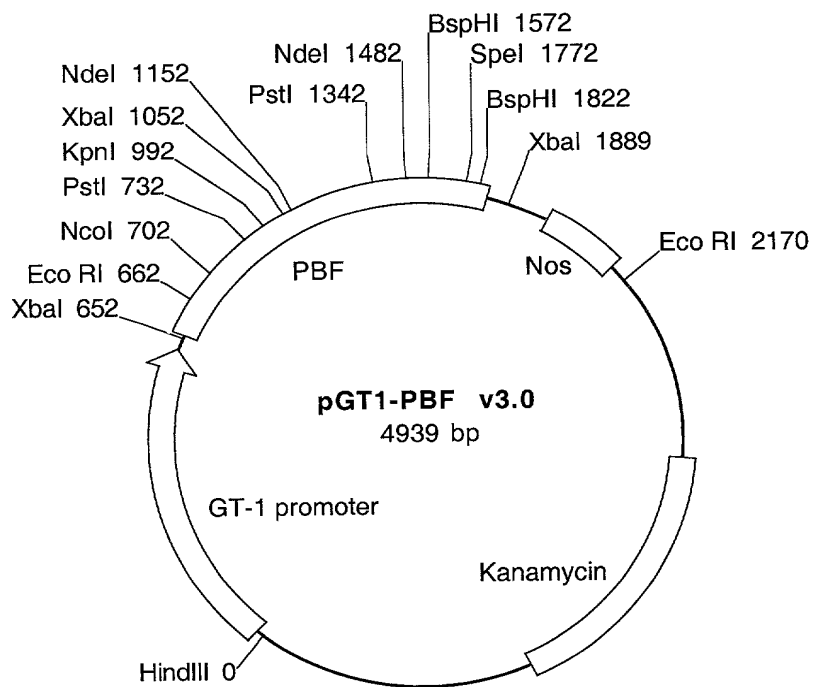


Fig. 11B

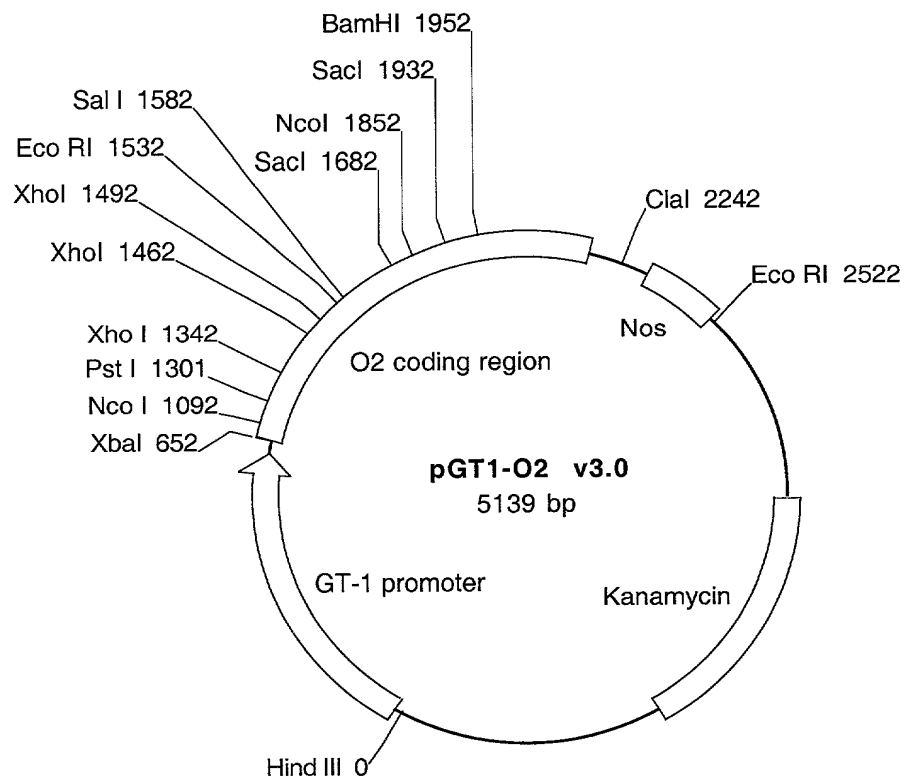


Fig. 11C

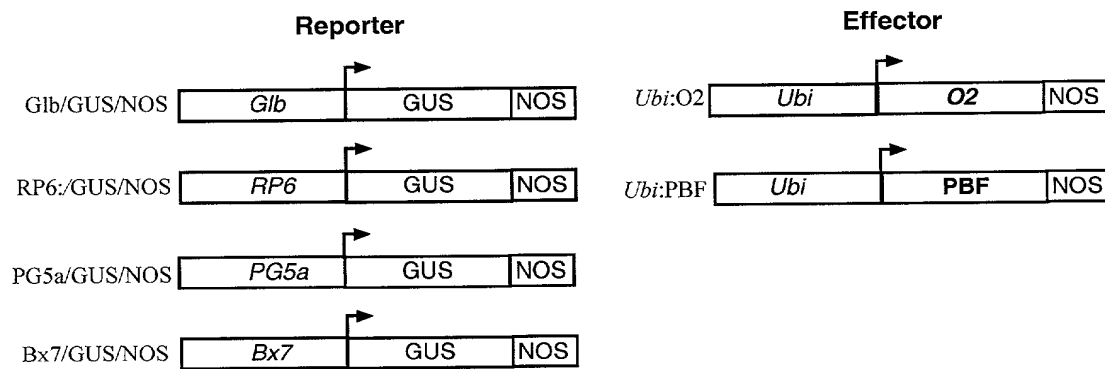


Fig. 12A

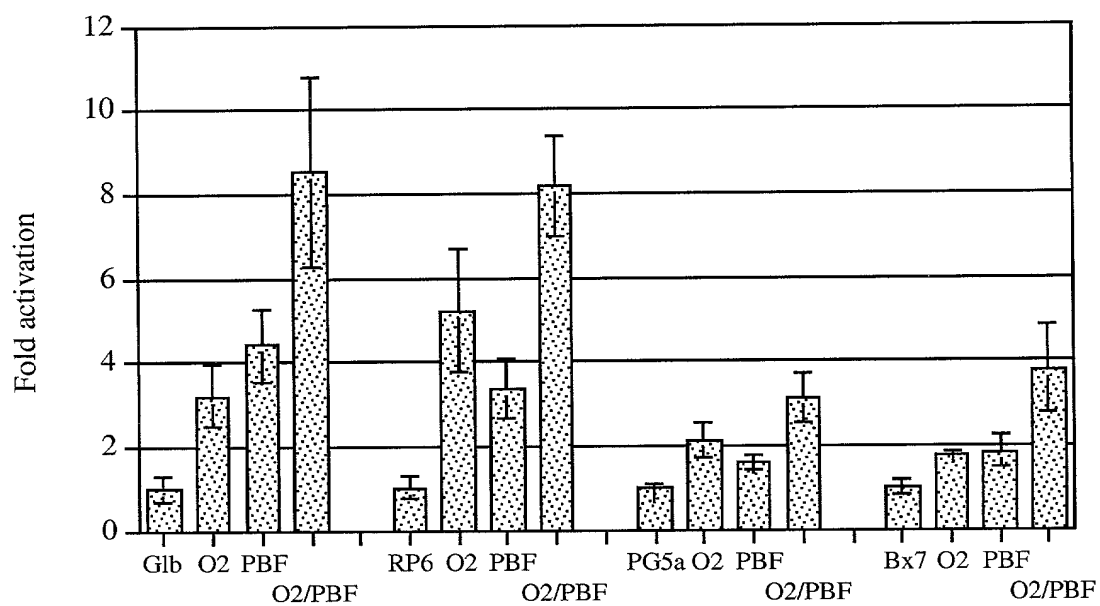


Fig. 12B

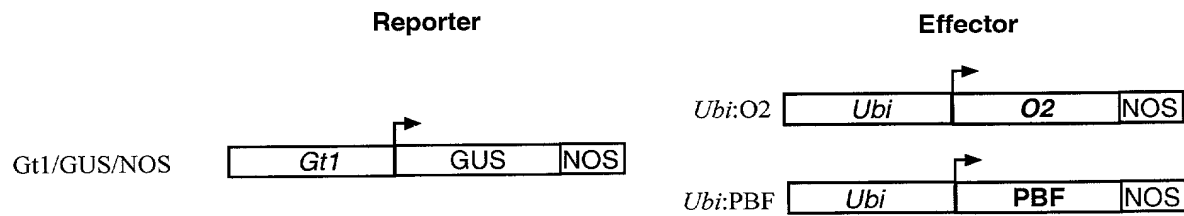


Fig. 13A

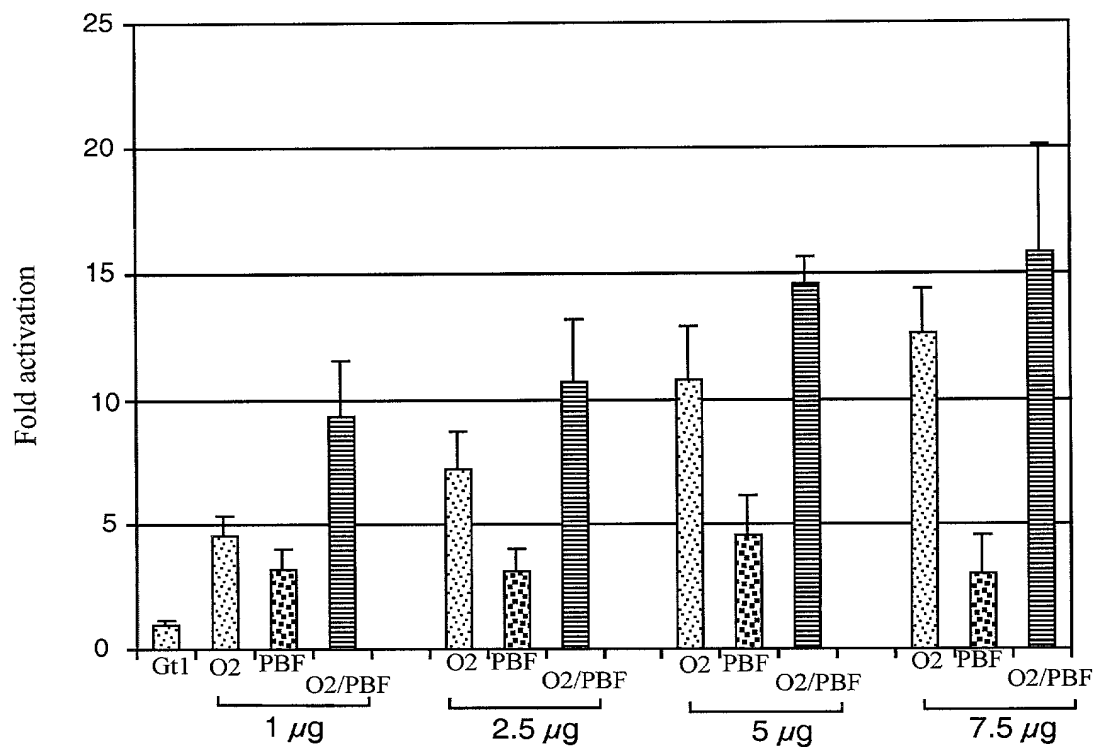


Fig. 13B

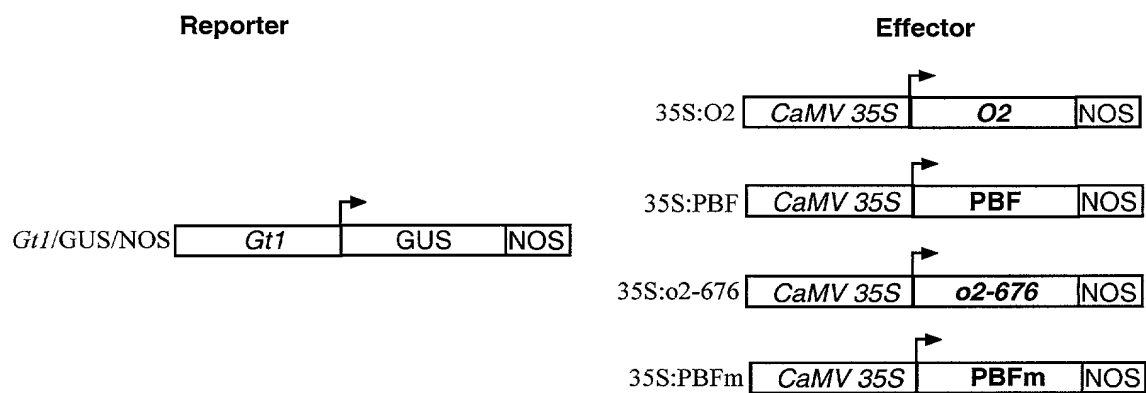


Fig. 14A

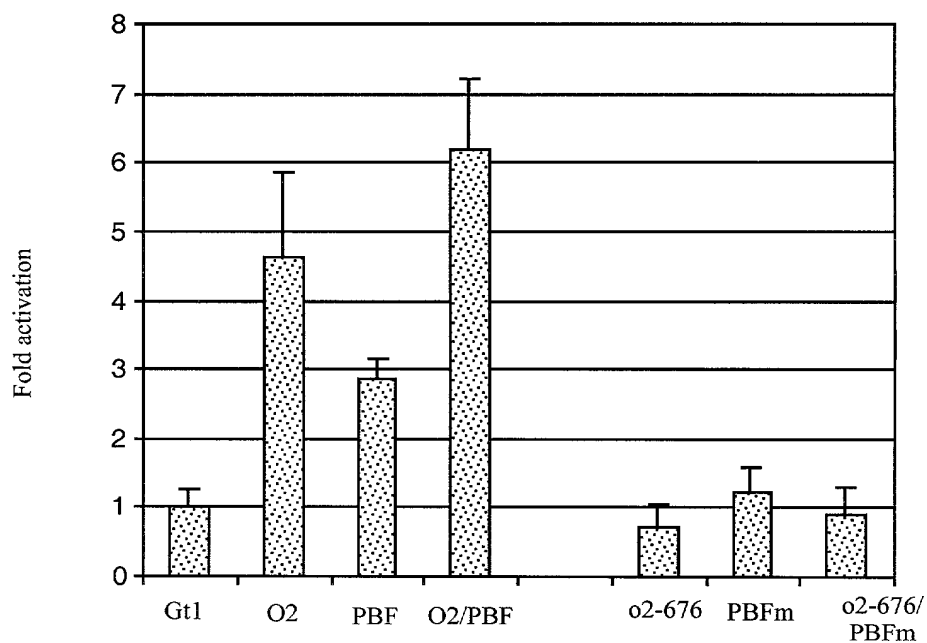


Fig. 14B

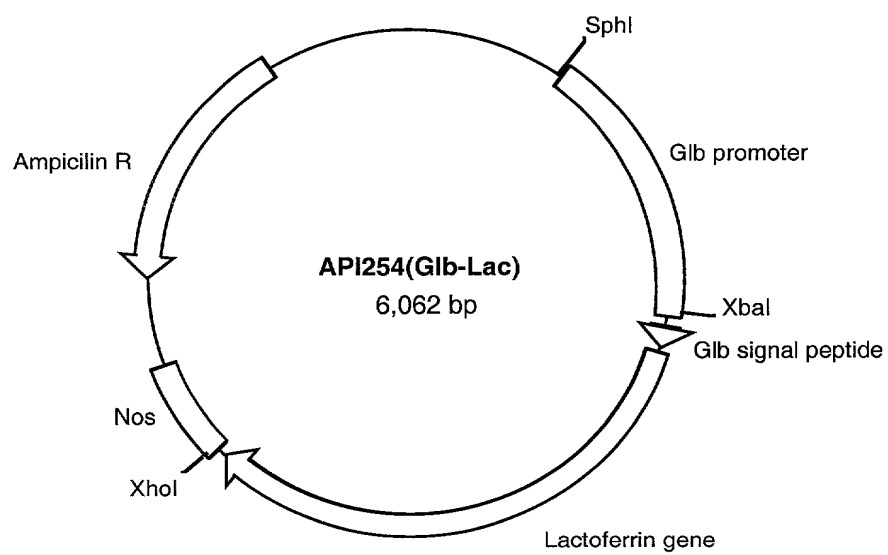


Fig. 15A

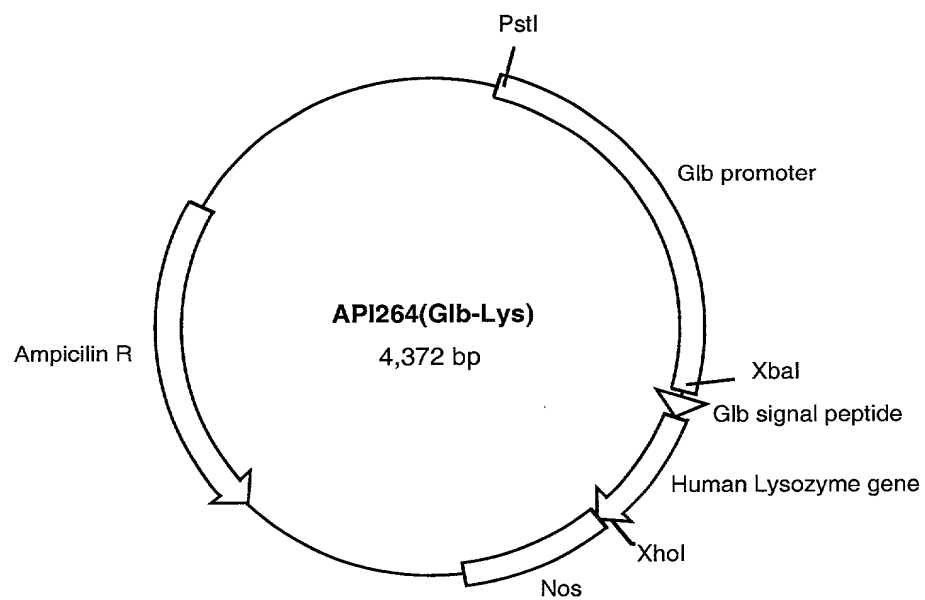


Fig. 15B

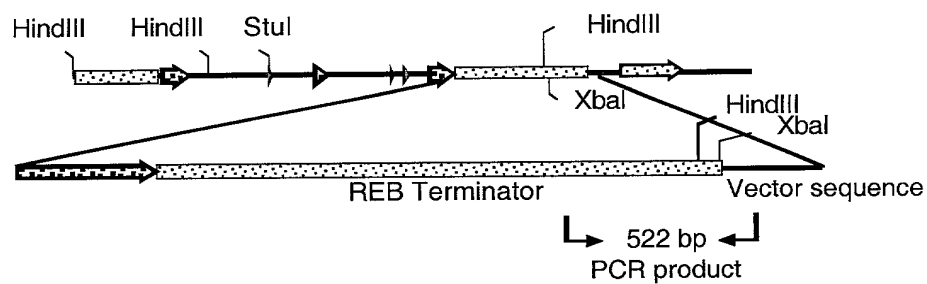


Fig. 16A

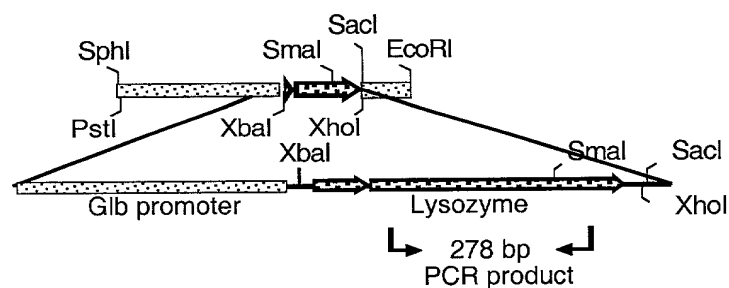


Fig. 16B

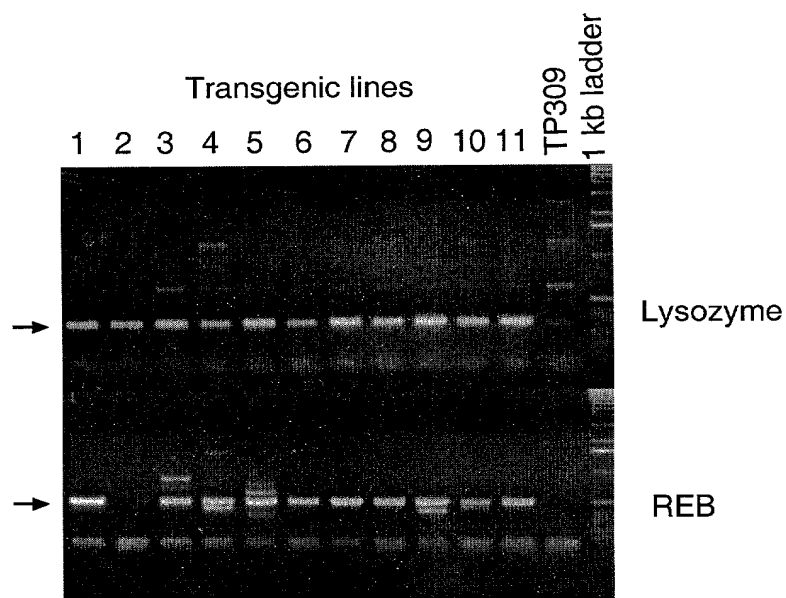


Fig. 16C

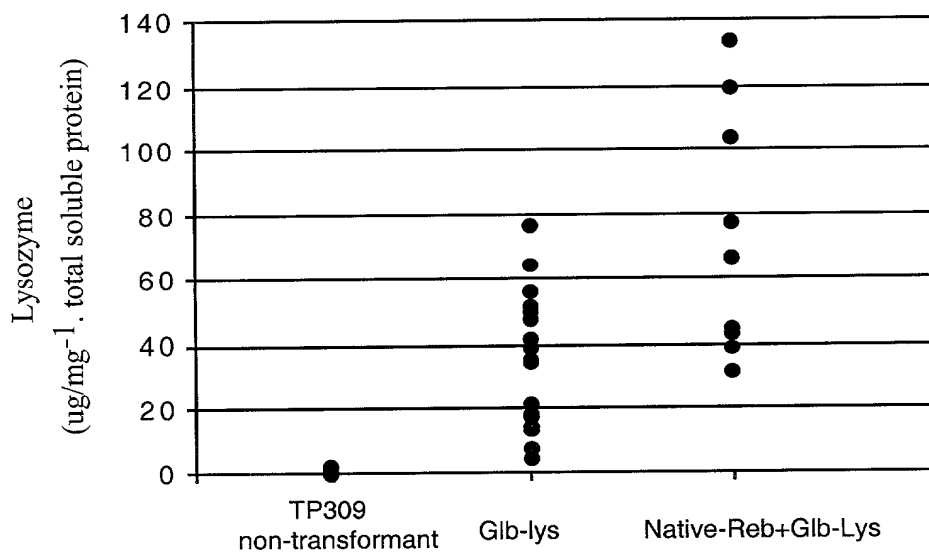


Fig. 17